

Guidelines for an Ocean Project or Programme Data Management Plan

Guidelines for an Ocean Project or Programme Data Management Plan



UNESCO 2025

This document emphasizes the importance of data management planning in activities which collect, create or otherwise acquire ocean data, ensuring data are preserved and accessible for analysis and to inform and underpin evidence and decision making. A Data Management Plan is a document outlining how to handle data throughout its lifecycle, ensuring proper collection, documentation, accessibility, and preservation. Key topics to be considered in a Data Management Plan include data collection and categorization, storage and processing, analysis, publication and access, archival, reuse, and erasure.

Additional benefits of a Data Management Plan include improving data maturity and literacy, effective storage and archival, efficient project or programme delivery and improved compliance with data sharing, contractual and licensing arrangements.

A detailed Data Management Plan Template is provided, covering the entire data lifecycle and guiding consideration of any ethical and legal issues around the data which may arise at any or all stages of the data lifecycle. Each section contains multiple subsections with specific questions to guide the development of a comprehensive Data Management Plan.

Prepared by the Decade Coordination Office for Ocean Data Sharing.

Authors:

Leadbetter, A.; Ariasis, F.; Buttigieg, P.-L.; Giorgetti, A.; Hanley, L.; Hebden, M.; Iona, A.; Lear, D.; Lopez, M.; Moulton, C.; Tyberghein, L.

For bibliographic purposes this document should be cited as follows:

UNESCO-IOC (2025). *Guidelines for an Ocean Project or Programme Data Management Plan*. Paris. Intergovernmental Oceanographic Commission of UNESCO. (IOC Manuals and Guides, 99) (English)

This publication supersedes IOC Manuals and Guides, [73](#) (2016): *Guidelines for a data management plan*

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.

TABLE OF CONTENTS

1. **CONTEXT 1**

2. **WHO IS THIS DOCUMENT FOR? 1**

3. **WHAT IS A DATA MANAGEMENT PLAN? 1**

4. **ADDITIONAL BENEFITS OF A DATA MANAGEMENT PLAN..... 3**

5. **REFERENCES 3**

6. **DATA MANAGEMENT PLAN TEMPLATE 4**

ANNEX: [DATA ROLES](#)

1. CONTEXT

New technologies, applications, and national commitments are revolutionising ocean data collection and access, providing a unique opportunity to gather and analyse data about our environment, the human use of marine resources, and create opportunities for better science and decision-making (Trice et al., 2021). Collaboration across activities such as observations, metadata and data assembly, quality assurance and control, and data publication is essential for effective data management, and supporting delivery of goals from the FAIR principles on data reusability.

Data management plays a key role in ocean observing systems, ensuring that data is not only collected but also preserved and made accessible for both present and long-term analysis and application (Tanhua et al., 2019). In this document we expect Data Management Plans to be created for projects (time-bound activities) or programmes (non-time bound activities) which collect, create or otherwise acquire data relating to the oceans. Clear roles and responsibilities of all project/programme and project/programme support team members should be determined at the start of a project/programme, ideally before data acquisition or data collection begins, especially when data activities are spread across multiple teams. This ensures that everyone understands their role in data management, and these responsibilities should be documented within the data management plan (UNESCO, 2016).

This document was prepared in response to a survey of Actions under the UN Decade of Ocean Science for Sustainable Development conducted in June 2024 by the Decade Coordination Office for Ocean Data Sharing. This survey identified that many Actions were limited in their ability to share data into regional or global systems as they lacked guidance in how to plan for data management, including data sharing and publication. This document does not provide a technical data management manual and should be used in conjunction with other documents including, but not limited to, the OceanOps Tutorial¹, the Ocean Biodiversity Information System Manual² and the Ocean Data Information System book³. It may also be supplemented by funder requirements for data management plans such as the Belmont Forum's Data and Digital Management Annex⁴.

2. WHO IS THIS DOCUMENT FOR?

This document is aimed at those who are responsible for the good data management of the outputs of a project or programme which creates, collects or otherwise acquires data relating to the oceans. This document may also be used by project or programme leads who are responsible for ensuring a data management plan is created in order to fulfill funding requirements.

3. WHAT IS A DATA MANAGEMENT PLAN?

A data management plan (DMP) is a document that outlines how to handle data collected or generated during a project or programme, and a structure for its ongoing management throughout the data life-cycle (see Figure 1). Its goal is to ensure that data is properly collected, documented, made accessible, and preserved for future use, as part of best practice, as well

¹ <https://www.ocean-ops.org/board>

² <https://manual.obis.org/>

³ <https://book.odis.org/>

⁴ <https://www.belmontforum.org/archives/resources/data-and-digital-outputs-management-annex-full>

as often being a requirement of funding and research organisations (UNESCO, 2016). DMPs support team members in considering all relevant aspects of data management from the inception of a project or programme, and invite team members to consider optimal handling, organising, documenting, and storing of data. DMPs also provide early recognition of the needs for the data to be met as a key result of the project delivery, rather than as a separate, retrospective, activity.

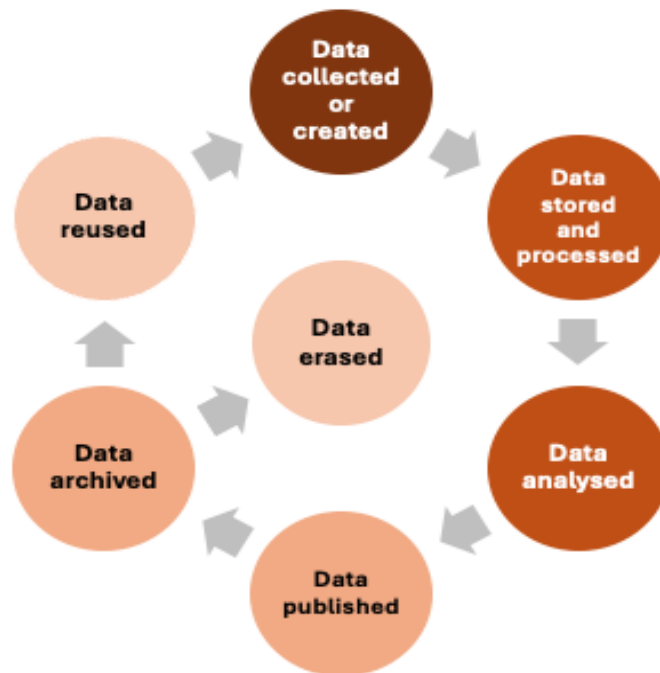


Figure 1. *An overview of the data lifecycle assumed in this document (after Leadbetter et al, 2019).*

A DMP should include a variety of questions around the following key topics to be considered across the different phases of the project and data lifecycle :

- data collected or created and its thematic categorisation
- data storage and processing
- data analysis, including data quality
- data publication and access
- data archival
- data reuse (sharing and publication)
- data erasure or deletion

In the planning stage of any project, it is important to involve specialist data managers as early as possible, and have discussions around the data collection requirements, the instruments and platforms to be used, the frequency of data collection/availability, data storage and accessibility requirements. Clear processing timelines, data availability policies, and metadata standards should be identified.

Once the necessary information is gathered, a data management team should be established and must draft a comprehensive plan outlining data assembly, processing, delivery timelines, and consideration of potential restrictions. A data management plan is also useful for identifying who is responsible for any data identified across its lifecycle. This plan should be reviewed and agreed upon by both the data management team and the project organisers before the project begins (UNESCO, 2016).

When ready, the data management plan should be easily accessible by the public. The usual practice is to place this document on the project website, redacting any sensitive information such as local network storage locations which may pose an IT security risk.

Funding agencies may have their own requirements for a data management plan and their own assessment criteria, such as the Belmont Forum's Data Management Plan Scorecard (Bishop et al., 2019), which should be taken into consideration alongside the guidance presented in this document.

4. ADDITIONAL BENEFITS OF A DATA MANAGEMENT PLAN

Other benefits and opportunities for including a DMP in a project, include:

- **Data Maturity** – improving data 'journeys' and increasing team capability in using data.
- **Data Literacy** – awareness raising, capacity building and breaking technical language barriers to support research project delivery.
- **Effective data storage, transfer and archival decisions** – reducing monetary and carbon costs and improving compliance with contractual and licensing arrangements and requirements.
- **Guidance on approach**, lessons learnt across projects with avoidance of duplication of effort.
- **Efficient project delivery** - avoidance of effort duplication, improved scheduling, team buy-in.
- **Funders** can gain **oversight** of trends across their funded projects, which can inform future funding opportunities.

This document supersedes previous guidance from the Intergovernmental Oceanographic Commission on Research Data Management Planning (UNESCO, 2016) to consider developments such as the Ocean Data Information System (ODIS; <https://odis.org/>) and the UN Decade of the Ocean for Sustainable Development's Data and Information Strategy, with its associated Implementation Plan.

Section 6 gives an overall description on the components and preparation of a data management plan (DMP).

5. REFERENCES

Bishop, B. W., Ungvari, J., Davis, R. I., Lee, T., Goudeseune, L., Virapongse, A., & Samors, R. J. (2019). *Belmont Forum Data Management Plan Scorecard*.
<https://doi.org/10.5281/zenodo.3530933>

- Gordon K (2013) *Principles of data management: facilitating information sharing second edition*. BCS
- Leadbetter, A., Carr, R., Flynn, S., Meaney, W., Moran, S., Bogan, Y., & Thomas, R. (2019). *Implementation of a Data Management Quality Management Framework at the Marine Institute, Ireland. Earth Science Informatics*, 1-13.
- Plotkin D. (2013) *Data stewardship: an actionable guide to effective data management and data governance*. Newnes
- Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., De Bruin, T., & Zhao, Z. (2019). *Ocean FAIR data services. Frontiers in Marine Science*, 6, 440.
- Trice, A., Robbins C., Philip, N. and Rumsey, M. (2021) *Challenges and Opportunities for Ocean Data to Advance Conservation and Management*, Ocean Conservancy, Washington D.C.
- UNESCO-IOC (2016). *Guidelines for a Research Data Management Plan*. Paris. Intergovernmental Oceanographic Commission of UNESCO, 16 pp. (*IOC Manuals and Guides*, 73) (English.) ([IOC/2016/MG/73](#)) Superseded by this document.
- UNESCO-IOC (2023). *Ocean Decade Data & Information Strategy*. Paris. Intergovernmental Oceanographic Commission of UNESCO, 35 pp. (*The Ocean Decade Series*, 45). <https://unesdoc.unesco.org/ark:/48223/pf0000386684.locale=en>

6. DATA MANAGEMENT PLAN TEMPLATE

Section 1. Data Management Plan Information
<i>Use this section to describe the history of the Data Management Plan.</i>
 <i>Who created it, who modified it and when. Alternative approaches to this may include simple versioning of data management plans.</i>
1.1 Data Management Plan prepared by
<i>The name and affiliation of the person (or people) preparing the Data Management Plans. Organisation contact e-mail addresses should be included.</i>
1.2 Data Management Plan publication date
<i>The date on which the Data Management Plan was first created.</i>
1.3 Data Management Plan history of changes
<i>The date of the most recent update of the Data Management Plan, or a version number.</i>

Section 2. Programme / Framework / Project Information

Use this section to describe the initiative collecting the data (where applicable).

2.1 Programme / Framework / Project number

Insert project reference number, if applicable. For example, the reference number given to the project by funders.

2.2 Programme / Framework / Project acronym

Insert the short name or acronym used to commonly identify the project creating or acquiring the data.

2.3 Programme / Framework / Project title

Insert the full name of the project creating or acquiring the data.

Section 3. Data Creation, Collection and/or Acquisition

This section is used to provide an overview of the data creation.

3.1 What data will be collected, created and/or acquired?

What parameters or attributes are being collected or created in these data, e.g. temperature, salinity, abundance of Mytilus Edulis, modelled temperature or salinity. Are these new data, or are the data re-used from existing resources?

Will any third party data be acquired? If so, from where and under what terms and conditions?

3.2 What methods will be used in data collection or creation?

What methods or best practices are used in the initial data collection or creation. Where appropriate, provide references to journal papers or links to documentation, such as from the Ocean Best Practices repository.

Also describe any platforms (such as research vessels or models) or instruments used in the data collection and how the data will be transferred from the platforms to the storage location.

3.3 What file formats will the data be collected, created or acquired in?

Please list the file formats used to store the data (e.g. Microsoft Excel, netCDF, OceanDataView, SQL Database, TIFF, MP4).

If possible, please indicate the expected volume of data to be produced (for example 10Mb; 1Tb etc...).

Section 3. Data Creation, Collection and/or Acquisition
3.4 What conventions or standards, if any, are used at the time of data collection, creation or acquisition?
<i>The preparation of data files often follows conventions or standards such as the Climate and Forecast (CF) conventions for NetCDF file formats, SeaDataNet for physical and biogeochemical data, including litter, or Darwin Core for biodiversity data.</i>
<i>Please list the conventions or standards used in preparation of the data files here.</i>
3.5 What ontology/vocabulary, if any, is used to name the parameters in the data files from the time of data creation, collection and/or acquisition?
<i>What ontology or vocabulary will be used for the data parameters?</i>
3.6 What is the backup strategy used to ensure the data files from this stage of the data lifecycle can be recovered in the event of an IT disaster?
<i>Please describe how continued access to the data will be ensured if there is an IT disaster, such as a security breach or corrupted local drive?</i>
3.7 (To be completed after data collection) What differences or discrepancies occurred between planning and defining the objective of the data creation (see Q. 2.2) or collection and the completion of data creation or collection?
<i>Differences or discrepancies may occur for operational reasons (such as a change in survey plans) or for other valid reasons. Giving details here may help later assessments of the appropriateness or completeness of the data for further analysis.</i>
3.8 Following data collection, creation or acquisition but before processing and storage, who may the data be shared with, under what conditions and when?
<i>At this stage of the data lifecycle, with whom may the data be shared (for example a project team or across an institution)? What restrictions are placed on any users at this stage of the data lifecycle (e.g. not for publication)? How soon after data creation, collection or acquisition may the potential users listed gain access to the data?</i>
3.9 How is the data creation, collection and/or acquisition to be funded?
<i>How is this stage of the data lifecycle funded?</i>
3.10 People, Institutions, Contact Details, Roles involved in Data Creation, Collection and/or Acquisition

Section 3. Data Creation, Collection and/or Acquisition

List the names, organisational affiliations, contact details and roles of all associated with the data.

Roles are outlined in Annexe 1.

Section 4. Data Storage and Processing

Use this section to describe any post-processing following the initial acquisition/collection detailed in Section 3.

4.1 What methods will be used in data processing?

What methods or best practices are used in the post-processing of the data. Where appropriate, provide references to journal papers or links to documentation, such as from the Ocean Best Practices repository.

4.2 What file formats will the data be in (or available in) once processing is complete?

Please list the file formats used to store the data (e.g. plain text formats, Microsoft Excel, netCDF, JSON, Application Programming Interface, Open Geospatial Consortium Web Services, Raster or Vector Geospatial Formats, OceanDataView, SQL Database, TIFF, MP4).

If possible, please indicate the expected volume of data to be produced (for example 10Mb; 1Tb etc...).

How often will the data be updated? For example, is the data static, or does it receive real-time or near-real time updates and on what frequency?

4.3 What conventions or standards, if any, are used in the preparation of the data file(s) at this stage of the data lifecycle?

The preparation of data files often follows conventions or standards such as the Climate and Forecast (CF) conventions for NetCDF file formats, SeaDataNet for physical and biogeochemical data, including litter, or Darwin Core for biodiversity data.

Section 4. Data Storage and Processing

Please list the conventions or standards used in preparation of the data files here.

4.4 What ontology/vocabulary, if any, is used to name the parameters in the data files at this stage of the data lifecycle?

What ontology or vocabulary will be used for the data parameters?

4.5 How will the quality of the data be assessed and reported?

What methods or best practices are used in assessing or quantifying the quality of the data. How will any errors or outliers in the data be identified and reported? How will any missing data values be identified and reported? Do these methods for data quality assessment and reporting meet the ends of the expected end users of the data.

Where appropriate, provide references to journal papers or links to documentation, such as from the Ocean Best Practices repository. Please describe any data flagging schemes in use, such as the [Argo data quality flags](#).

Please detail how you will record both in the data, accompanying metadata (data about the data) and documentation what quality control steps have been taken and who (person, role and/or organisation) will undertake the quality control.

4.6 What is the backup strategy used to ensure the data files from this stage of the data lifecycle can be recovered in the event of an IT disaster?

Please describe how continued access to the data will be ensured if there is an IT disaster, such as a security breach or corrupted local drive?

4.7 Following data processing and storage but before data analysis, who may the data be shared with, under what conditions and when?

At this stage of the data lifecycle, with whom may the data be shared (for example a project team or across an institution)? What restrictions are placed on any users at this stage of the data lifecycle (e.g. not for publication)? How soon after data processing and storage may the potential users listed gain access to the data?

4.8 How will the data storage and processing be funded?

How is this stage of the data lifecycle to be paid for?

Section 4. Data Storage and Processing

4.9 People, Institutions, Contact Details, Roles involved in Data Storage and Processing

List the names, organisational affiliations, contact details and roles of all associated with the data.

Roles are outlined in Annexe 1.

Section 5. Data Analysis

Use this section to describe any more detailed analysis which will take place, after post-processing, in order to create a data product.

5.1 What methods will be used in data analysis?

What methods or best practices are used in the analysis of the data. Where appropriate, provide references to journal papers or links to documentation, such as from the Ocean Best Practices repository.

5.2 What file formats will the data be in once analysis is complete?

Please list the file formats used to store the data (e.g. plain text formats, Microsoft Excel, netCDF, JSON, Application Programming Interface, Open Geospatial Consortium Web Services, Raster or Vector Geospatial Formats, OceanDataView, SQL Database, TIFF, MP4).

If possible, please indicate the expected volume of data to be produced (for example 10Mb; 1Tb etc...).

5.3 What conventions or standards, if any, are used in the preparation of the data file(s) at this stage of the data lifecycle?

The preparation of data files often follows conventions or standards such as the Climate and Forecast (CF) conventions for NetCDF file formats, SeaDataNet for physical and biogeochemical data, including litter, or Darwin Core for biodiversity data.

Please list the conventions or standards used in preparation of the data files here.

5.4 What ontology/vocabulary, if any, is used to name the parameters in the data files at this stage of the data lifecycle?

What ontology or vocabulary will be used for the data parameters?

Section 5. Data Analysis

5.5 How will any further assessments of the quality of the data be undertaken and reported?

What methods or best practices are used in assessing or quantifying the quality of the data. How will any errors or outliers in the data be identified and reported? How will any missing data values be identified and reported? Do these methods for data quality assessment and reporting meet the ends of the expected end users of the data.

Where appropriate, provide references to journal papers or links to documentation, such as from the Ocean Best Practices repository. Please describe any data flagging schemes in use, such as the [Argo data quality flags](#).

Please detail how you will record both in the data, accompanying metadata (data about the data) and documentation what quality control steps have been taken and who (person, role and/or organisation) will undertake the quality control.

5.6 What is the backup strategy used to ensure the data files from this stage of the data lifecycle can be recovered in the event of an IT disaster?

Please describe how continued access to the data will be ensured if there is an IT disaster, such as a security breach or corrupted local drive?

5.7 Following data analysis but before data publication, who may the data be shared with and under what conditions?

At this stage of the data lifecycle, with whom may the data be shared (for example a project team or across an institution)? What restrictions are placed on any users at this stage of the data lifecycle (e.g. not for publication)? How soon after data analysis may the potential users listed gain access to the data?

5.8 How is the Data Analysis to be funded?

How is this stage of the data lifecycle to be paid for?

5.9 People, Institutions, Contact Details, Roles involved in Data Analysis

List the names, organisational affiliations, contact details and roles of all associated with the data.

Roles are outlined in Annexe 1.

Section 6. Data Publication

This section is used to provide information about the metadata (data about the data) which will be used to aid in discoverability (findability) and re-use of the data. Also described are the services used to publish the data for access by other users.

Data publication in this sense may be an informal process, with data potentially being made available from institutional services and is not a guarantee of the long-term preservation or archival of the data in a specialist data facility. However, this lifecycle phase may also be completed by the long-term repository responsible for archival.

This lifecycle phase does not require the formal publication of data with a Digital Object Identifier, which is considered under Data Archival below.

6.1 What metadata (data about the data) standards to describe the data are used by either the research community or the funding agencies?

Metadata standards are often set at a national or international level. Often the metadata standards are based on a profile of a broader standard (such as the EU's INSPIRE Spatial Data Infrastructure and the US' Earth Observing System Data and Information System based on ISO19115; or OBIS' use of the Ecological Metadata Language).

These metadata standards will often describe the way in which the data are geo-referenced, so please describe how your data will be located in space and time.

6.2 What keywords will you use to describe the data, and which vocabulary will you take them from?

Keywords support the findability or discovery of a data. Metadata standards often describe which vocabulary (set of keywords) to use in describing a data (such as from the NERC Vocabulary Server, the CF Standard Names list or the Environment Ontology (EnvO)).

Please list here the vocabularies and keywords you will use in describing the data.

6.3 What other documentation can be provided to support the data?

Are there any other documents which will provide additional contextual information supporting the data? These may include but are not limited to research vessel survey reports or data description papers.

Section 6. Data Publication
<i>How will the processing of the data (who did what to the data, when and why) from collection to final data product be described?</i>
6.4 How will the metadata be published to make it generally available?
<i>How will the metadata created for the data be published and made generally, publicly available? Which catalogue or system will be used to allow others to discover the existence of the data?</i>
6.5 What services are used to publish the data?
<i>Please list any web services which are used to make the underlying data available (not the metadata), such as Web Map Services, Esri ArcGIS Online, Erddap data servers, cloud storage (such as Amazon S3 or Microsoft Azure object storage).</i>
6.6 Does the data require any further transformation (file formats, attributes, conventions) in order to be published? If yes, please describe.
<i>Please list any further changes in file formats, the naming of attributes or alterations to the conventions used in the preparation of the data in order to allow the Data Publication lifecycle phase to be achieved.</i>
6.7 What licence will the data be published under?
<i>An open license (for example a Creative Commons or Open Government Licence) is recommended for data sharing, especially for alignment with the FAIR principles. However, this is not always possible.</i>
<i>Please provide a link to the license under which the data will be shared.</i>
6.8 Who may the data be shared with or published to?
<i>Open sharing of the data under an open licence is recommended, but if this is not possible (for example the data contains sensitive or personal data) please list the individuals or organisations with whom the data may be shared and under what arrangements or conditions (e.g. formal data sharing agreements, contractual agreements, etc...).</i>
6.9 Are there any embargo periods to be placed on the data sharing before external publication?
<i>Is there a time during which the data will have been collected and processed but which should not be shared wider? If so, please provide the details and an explanation of why this embargo period is in place here.</i>
6.10 When the data are shared or published, are there any uses for which the data are not suitable?
<i>For example, some bathymetry data may not be suitable for navigation.</i>

Section 6. Data Publication

6.11 Are there any further terms of use to be placed on the data?

Please list any further terms of use that are to be placed on the data, for example for warnings or flood forecasts it may be appropriate to add a caveat such as “Users may not redistribute the material if has been modified” or “material content of the warnings may not be altered, and the warnings be kept up to date”

6.12 How is Data Publication to be funded?

How is this stage of the data lifecycle to be paid for?

6.13 People, Institutions, Contact Details, Roles involved in Data Publication

List the names, organisational affiliations, contact details and roles of all associated with the data.

Roles are outlined in Annexe 2.

Section 7. Data Archival

Use this section to describe the long-term archival of the data, normally in a long-term data archival, data assembly centre or data repository facility.

7.1 Which parts of the data will be selected for long-term archival / preservation?

If there are parts of the data which will not be selected for long-term preservation or archival, such as part-processed data, please provide a description and a reason here.

How will you make the version history of the data (from raw data to harmonised data to fully processed data) transparent to end users and make data from any stage of this process available to end users?

7.2 What is the long-term importance or value of the data?

Please give a description of the anticipated long-term value of the data, for example will the data be used in sustainable growth or sustainable development, health or food security?

7.3 What is the long-term preservation plan for the data?

Please describe which archive, data facility or data assembly centre the data will be submitted to for long-term preservation and indicate if the submission of the data has been discussed with them.

Section 7. Data Archival

The archive, data facility or data assembly centre should be encouraged to publish the data's metadata to the Ocean Data and Information System (ODIS) and to openly publish the data (where applicable) using web standards.

Any data aggregators that the archive, data facility or data assembly centre is connected to and which may validate the data and metadata should also be documented.

Are there any other intended archival or preservation approaches, such as formal data publication with a digital object identifier (doi)?

7.4 Does the data require any further transformation (file formats, attributes, conventions) or quality control in order to be archived? If yes, please describe.

Please list any further changes in file formats, the naming of attributes or alterations to the conventions used in the preparation of the data in order to allow the Data Archival lifecycle phase to be achieved. For example, is there anything further required to make the data meet the FAIR principles, or are geometry validations required?

7.5 Will the data license, terms and conditions or any other restrictions on use change between Data Publication and Data Archival? If so, please indicate how.

Questions 6.7-6.11 detail various licensing and restrictions on the data. Please indicate any changes to the answers to those questions resulting from the Data Archival phase of the lifecycle.

7.6 How is Data Archival to be funded?

How is this stage of the data lifecycle to be funded?

7.7 People, Institutions, Contact Details, Roles involved in Data Archival

List the names, organisational affiliations, contact details and roles of all associated with the data.

Roles are outlined in Annexe 1.

Section 8. Data reuse

Are there any further documents required to enable reuse of the data.

Are there any metrics in place to monitor how and where the data are reused.

8.1 What other documentation can be provided, or linked to, to support the data?

Are there any other documents which will provide additional contextual information supporting the data? These may include but are not limited to research vessel survey reports or data description papers.

How will the processing of the data (who did what to the data, when and why) from collection to final data product be described?

8.2 What metrics are in place to monitor the reuse of the data and its impact?

How will the reuse of the data be tracked? For example through citations of a digital object identifier for the dataset.

Who will be responsible for maintaining and reporting any chosen metrics and how will they be funded?

Section 9. Data erasure

Under what circumstances may the data be deleted. What are the implications of this and how will users be notified?

9.1 What foreseeable circumstances, excluding hardware failure or cybersecurity issues, may trigger the intended complete erasure of the data?

Are there intended circumstances in which the data may be erased. This could be a superseding of model data with a new analysis, or for data containing personally identifiable information a request for a data subject.

Please do not include unintended circumstances here, as they should be covered elsewhere by data backup questions.

9.2 What would be the consequences of erasing the data?

What would be the consequences to any users of the data if the intended erasure of the data was undertaken.

Section 9. Data erasure 9.3 How would users of the data be informed of an intended erasure of the data? <i>What steps would be taken to give users warning of the erasure of the data? How long a period of notice or warning of the intended erasure would be given? What alternative sources of similar data exist which users could switch to?</i>
9.4 People, Institutions, Contact Details, Roles involved in Data Erasure <i>List the names, organisational affiliations, contact details and roles of all associated with the data.</i> <i>Roles are outlined in Annexe 1.</i>

The questions below on Ethical and Legal compliance may apply at any or all stages of the data lifecycle, and the answers to the questions may change across the lifecycle. They are included here in a separate section in order to keep the above sections as short as possible.

Section A. Ethical and Legal Compliance <i>This section is used to describe any legal or ethical issues which may be associated with the data, including Intellectual Property Rights.</i>
<i>The CARE Principles for Indigenous Data Governance, where applicable, are considered in this section.</i>
A.1 Is there an ethics policy which was required to be followed during the creation or processing and analysis of the data? <i>If yes, please provide a link to the policy and explain how the data meets the requirements of that policy.</i>
A.2 Is there a diversity policy which was required to be followed during the creation or processing and analysis of the data? <i>If yes, please provide a link to the policy and explain how the data meets the requirements of that policy.</i>
A.3 Are there any considerations related to Collective Benefit for indigenous peoples associated with the data? <i>The CARE Principles consider the following topics under Collective Benefit. Please address these here if applicable to the data:</i> <i>Inclusive development and innovation</i> <i>Improved governance and citizen engagement</i> <i>Equitable outcomes</i>

Section A. Ethical and Legal Compliance
A.4 Are there any considerations related to the Authority of indigenous peoples to Control the data?
<i>The CARE Principles consider the following topics under Authority to Control. Please address these here if applicable to the data:</i> • <i>Recognising rights and interests</i> • <i>Data for governance</i> • <i>Governance of data</i>
A.5 Are there any considerations related to Responsibility towards supporting indigenous peoples' self-determination and collective benefit associated with the data?
<i>The CARE Principles consider the following topics under Responsibility. Please address these here if applicable to the data:</i> • <i>Positive relationships</i> • <i>Expanding capability and capacity</i> • <i>Indigenous languages and worldviews</i>
A.6 Are there any considerations related to the Ethics of indigenous peoples' rights and wellbeing associated with the data?
<i>The CARE Principles consider the following topics under Responsibility. Please address these here if applicable to the data:</i> • <i>For minimising harm and maximising benefit</i> • <i>For justice</i> • <i>For future use</i>
A.7 Are there any further ethical considerations associated with the data?
<i>Please list any further ethical considerations or issues associated with the data.</i>
A.8 Is there a data policy associated with the data describing access and intellectual property rights of the data?
<i>If yes, please provide a link to the data policy and explain how the data meets the requirements of that policy and what are the implications for Intellectual Property Rights associated with the data.</i>
A.9 Which individual or organisation owns the copyright of the data?
<i>List the name and contact details of the individuals or organisations who own the copyright and/or intellectual property rights associated with the data.</i>

ANNEX

DATA ROLES

The following data roles were described in Leadbetter *et al.*, 2019 and have been used in training courses on the IODE Data Quality Management Framework for National Oceanographic Data Centres:

Data owner

Following the definition in Gordon (2013), the role of Data Owner has the authority in the organisation to agree a dataset's classification and the retention schedule for a dataset. While the person tasked with this role may also be responsible for the stewardship or curation of the data values, the Data Owner is more likely to be in a team managerial role. A primary goal for this role is ensuring good data governance is achieved.

Data coordinator

The Data Coordinator role is closely aligned with the Data Administrator role as described in Gordon (2013). The Data Coordinator is responsible for the processes around data management within an organisational unit. Their responsibilities include oversight of the cataloguing of datasets whose Data Owner is a member of the business unit (in this model third-party datasets are catalogued by a centralised Data Management team); and facilitating the quality assurance of the data management processes in that unit. The Data Coordinator also acts as a liaison point between the central IT services in the organisation and their business unit which allows for data publication through centralised services. Through regular meetings between the Data Coordinators from the various business units, cross-organisational coordination of data management processes can be achieved.

Data steward

The Data Steward is involved with a dataset on a daily basis, and as such is responsible for many of the day-to-day activities around a dataset including the quality of the data; ensuring its safe archival and storage; and providing the required metadata and documentation around the dataset. Due to the technical scientific nature of their work within the organisational context, the Data Stewards will often blend aspects of the Business Data Steward and Technical Data Steward roles identified in Plotkin (2013). Therefore, as domain scientific experts they will understand the business needs fulfilled by the data they collect and curate but also will often have technical knowledge of database operations and numerical computing in scripting language environments. A Data Steward needs to operate within the bounds of organisational responsibilities and guidelines, and therefore needs to both be aware of these and other legislative requirements which may impact the datasets for which they are responsible; and to be supported by the organisation with appropriate training.

Data protection officer

A Data Protection Officer (DPO) is a leadership role in enterprise security (techniques and strategies for decreasing the risk of unauthorized access to data and IT systems and information). In Europe, this role is required by the General Data Protection Regulation (GDPR). DPOs are responsible for overseeing data protection strategy and implementation to ensure compliance with GDPR, or other legal requirements.

**Intergovernmental Oceanographic Commission
of UNESCO**

7 Place de Fontenoy
75352 Paris Cedex 07 SP, France
<http://ioc.unesco.org>

UNESCO/IOC Project Office for IODE

InnovOcean Campus
Jacobsenstraat 1
8400 Oostende, Belgium
Email: info@iode.org
<http://www.iode.org>